

# “Learning to be Learning Disabled:” Marie Clay’s Seminal Contribution to the Response to Intervention Approach to Identifying Specific Reading Disability

Frank R. Vellutino, Professor, The University at Albany, State University of New York

---

## Author’s Note

In this article, I provide a personal perspective regarding what I believe to be Marie Clay’s seminal contribution to the response to intervention (RTI) approach to identifying children who may be afflicted with organically based learning disability (specific reading disability) that puts them at risk for long-term reading difficulties. After providing historic context, I discuss, in some detail, an unsolicited article Marie Clay sent to me that led to a paradigm shift in my own reading research and stimulated several intervention studies I and my colleagues conducted that provided initial support for what I have come to call her *learning to be learning disabled* perspective. I discuss results from two of these studies and conclude with a brief discussion of concerns emanating from RTI models that emphasize assessment and identification rather than responsive instruction and prevention.

---

## Background

In 1988, Marie Clay and I were invited participants in a 2-day, weekend conference at the University of Virginia that was sponsored by scholars at the Curry School of Education. The conference has come to be an annual event that highlights the George Graham Lecture, which features distinguished scholars whose research is concerned with theory and practice in the study of literacy development. The lecture series was initiated in 1985 and, in its earliest iteration, was held three times a year.

I was the first George Graham lecturer and Marie Clay was the ninth. The proceedings usually begin on Friday and commence with an informal exchange of ideas between the featured speaker and other participants including faculty, other scholars, students at the univer-

sity, and invited guests from the educational community. The George Graham Lecture is presented on Saturday morning and, after a lunch break, a panel of scholars and educators discuss the morning lecture. Then the audience is given the opportunity to ask the featured speaker and the panelists any questions prompted by the panel discussion. I was an invited panelist who responded to Marie Clay’s lecture, along with the late Edmund Henderson, a distinguished scholar in the study of literacy development, and the late James Deese, a distinguished scholar in the study of learning, cognition, and psycholinguistics. The lecture was excellent; the questions were tough; and the interchange was productive and informative.

Given that the majority of the audience at this lecture consisted of educators and graduate students interested in literacy development and instruction (which typifies the participants in the George Graham Lectures), the lion’s share of the questions were concerned with the major components of Reading Recovery and the philosophy of instruction that shaped those components. Reading Recovery, as an early intervention, was relatively new in North America at the time, and because the approach to instruction advocated by scholars influenced by Edmund Henderson’s stage-based theory of reading and spelling development (Henderson, 1990) was appreciably different from the Reading Recovery approach, in terms of relative emphasis on “word study,” one could say that Marie Clay was in “enemy territory.” Nevertheless, she answered the questions honestly and forthrightly and presented the philosophic foundations of Reading Recovery in a clear, articulate, and compelling manner.

One aspect of Marie Clay’s lecture that to me was rather striking was her skepticism of learning disability (LD) as a viable psychological and educational construct. Also striking was her utter disdain of conventional psychometric approaches to the identification of LD in struggling readers which typically used the IQ-achievement discrep-

ancy along with commonly employed exclusionary factors (e.g., sensory deficits, emotional and behavioral problems, cultural disadvantage, etc.) as central defining criteria. These attitudes were also evident in the interactions with Marie during the more informal exchange of ideas at the meeting held on the day before the George Graham Lecture and during informal conversations I had with her discussing these ideas over cocktails and dinner. I found myself resonating to her perspective because my own research in the study of early reading difficulties and my clinical work with both struggling readers and children with learning difficulties in other areas made me at least as skeptical as she was about the concept of LD, in terms of its implementation in both research and educational practice.

I had long been concerned about the ill-defined nature of the concept of LD and the fact that it lacked strong scientific validation. Moreover, I was becoming increasingly alarmed that psychometric/exclusionary approaches to the identification of LD were unduly inflating the number of struggling readers being diagnosed as “disabled learners” in our public schools, owing to the use of imprecise and often invalid assessment techniques on the part of practitioners adopting such approaches. It was also clear to me, from the large number of “diagnostic” reports I had read, written by school psychologists, LD practitioners, and other clinical personnel, that conventional approaches to LD classification did little or nothing to inform instruction or remedial planning and typically led to low expectations of children assigned an LD label. I also felt that LD identification and classification practices were inherently discriminatory, given that most definitions of LD excluded children whose learning difficulties were associated with cultural disadvantage. Thus, I found little to disagree with in Marie Clay’s George Graham Lecture or in comments she made during informal discussions we had, at least as regards her views of conventional approaches to the diagnosis and treatment of reading and other learning difficulties.

Yet, Marie Clay may not have been as certain as I was that we had very similar views about these issues because several weeks after her George Graham Lecture, I received a letter from her expressing “some trepidation” in sending me a reprint of an article she wrote entitled “Learning to be Learning Disabled,” published the year before the lecture (Clay, 1987). The article provided a brilliant and compelling amplification of many of the concerns she expressed during the George Graham week-

end and I was greatly impressed by its contents — not only because they mirrored so many of my own views, but, more importantly, because they provided strong reason to believe that virtually all studies assessing the causes and correlates of reading and other learning difficulties were confounded by a major design problem, namely the failure to control for a child’s educational and personal history. My own research over the years has been motivated by my interest in evaluating the cognitive underpinnings of literacy and I, along with my colleagues and students, had already conducted a great deal

---

**It must be acknowledged that Marie Clay was actually the first reading researcher to use RTI to identify children who might be afflicted by organically based reading difficulties, although Reading Recovery, as originally conceived, was not designed for this purpose. Thus, her contribution to the RTI movement was seminal and certainly set the stage for subsequent intervention research that served to give this movement even greater momentum.**

---

of research assessing the viability of popular and influential theories of reading disability (e.g., perceptual deficit versus language-based theories). Our research, it seemed, had begun to distinguish between probable and improbable causes of reading difficulties, but there was always this nagging feeling that even our most promising results could not be readily generalized to all populations of struggling readers because our studies had not controlled for experiential and instructional factors that might have contributed to individual and group differences among children in our research samples. Indeed, this flaw was not only true of our studies, but was pervasive in the reading disability literature at large. Thus, as one might imagine, Marie Clay’s article had a sobering and influential effect on me because it convinced me that the external validity of our findings was as questionable as that of other studies in the reading disability literature.

Nevertheless, the arrival of her article was timely because I and my colleagues had been heavily involved in the process of preparing a grant proposal requesting funds for a large-scale intervention study designed to assess the utility of using an RTI approach to identifying children with specific reading disability in lieu of more conventional psychometric/exclusionary approaches. Marie's article provided us with both a strong rationale and a viable model for conducting reading intervention research that held promise for distinguishing between reading difficulties caused primarily by experiential and instructional limitations and reading difficulties caused primarily by basic cognitive limitations of constitutional origin. Fortunately, the study was funded and became the first in a series of intervention studies we conducted, each of which provided strong support for the utility of using an intervention-based approach to diagnostic assessment and equally strong support for the use of early intervention as a means of preventing long-term reading difficulties in children who might otherwise be classified as "reading disabled" (Scanlon, Gelzheiser, Vellutino, Schatschneider, & Sweeney, 2008; Vellutino, et al., 1996; Vellutino, Scanlon, Zhang, & Schatschneider, 2008). Results from these studies, among others (e.g., Torgesen, Rose, Lindamood, Conway, & Garvan, 1999) have become influential change agents in promoting the use of RTI approaches to identifying children afflicted by specific reading disability in lieu of more conventional psychometric/exclusionary approaches to accomplishing this objective, and, more importantly, in creating incentive for reducing the number of children classified as disabled learners by putting the focus on the need for more individualized, more responsive, and more effective classroom and remedial instruction.

However, Marie Clay's Reading Recovery, as a vehicle for identifying children at risk for long-term reading difficulties and thus in need of special educational services, predated these intervention studies. Therefore, it must be acknowledged that Marie Clay was actually the first reading researcher to use RTI to identify children who might be afflicted by organically based reading difficulties, although Reading Recovery, as originally conceived, was not designed for this purpose. Thus, her contribution to the RTI movement was seminal and certainly set the stage for subsequent intervention research that served to give this movement even greater momentum. Ironically, the seminal nature of her contribution has not been widely recognized. So, in the interest of promoting such recognition, I would like to share with the reader certain

of the ideas presented in Marie's 1987 article that were especially influential in motivating and shaping our own intervention research. I will then briefly discuss two of our intervention studies that provide strong support for her contention that most reading difficulties are caused primarily by experiential and instructional factors rather than by organically derived cognitive deficits. These studies also provide strong evidence in support of her contention that results from intelligence tests are not useful criteria for identifying struggling readers whose reading difficulties may be caused primarily by organically derived cognitive deficits.

### Learning to be Learning Disabled

Marie Clay's article is divided into three major sections and focuses throughout on difficulties in acquiring literacy skills as the prime example of learning difficulties. In Part I (Unresolved Problems), the article provides scholarly commentary and research evidence: (a) that specific learning disability is a "term that defied definition" and that existing definitions were based more on social and political expedience than on any compelling research documenting clear-cut differences between what she called "event-based causes" of reading and writing difficulties related to the child's personal and educational histories and organically based causes of such difficulties; (b) that assessment of potential for reading achievement and identification of specific reading disability based on intelligence test scores is misguided and discriminatory; and (c) that learning disabled and low-achieving readers are not distinguishable groups, at least not by virtue of conventional classification criteria. As regards the definitional problem, she cites research documenting the absence of any consensus, neither among researchers nor among practitioners regarding how to conceptualize or operationalize learning disability and underscored the folly and potential harm in legislating definitions of learning disability in the absence of such consensus (United States Office of Education, 1977). This is a point well made. Although we have made progress in distinguishing between probable and improbable causes of specific reading disability, as I indicated earlier, it seems fair to say that there is yet no definitive research identifying organically based causes of the disorder, although some of the research holds promise for doing so (see Vellutino, Fletcher, Snowling, & Scanlon, 2004 for a recent review).

In commenting on the use of intelligence tests to define reading disability, Marie Clay criticized this practice on

conceptual grounds and likened the ability to learn to read as akin to one's ability to acquire an intellectually low-level skill such as driving a motor vehicle insofar as both depend heavily on "self-scheduled" learning where the learner eventually becomes the main teacher through actively engaged experience and insofar as both generate a wide range of individual differences in ultimate levels of proficiency, despite even very low intelligence. I was struck by the following quote, which summarizes her perspective quite nicely:

I will assume that the brain's capacity to handle reading is somewhat like the brain's capacity to handle driving, and that some people become better drivers and some become better readers. The two activities have much in common. After an initial introduction, much of the learning is self-scheduled and the learner becomes the main teacher. Also, effective driving and effective reading have many analogous processing functions in the automatic visual processing area.

If the brain's capacity to handle reading is much like the brain's capacity to handle driving, then it does not make much sense to use intelligence as a criterion for reading attainment. I am then free to think of reading as going well above any intelligence predictions, or below. I could coach the child, taking him from where he is to somewhere else without even considering intelligence. I can readily accommodate the fact that five out of eight Downs Syndrome children recently retested by my colleague Kay Irwin, had chronological ages of 11, reading comprehension ages of seven or eight, and mental and language ages of around five years.

It is amazing that we have gone on using intelligence test scores as indicators of reading and writing potential for so long. They are fundamental to most definitions of LD which only include children with intelligence scores of average and above. This is grossly discriminatory if it determines who gets services. They are also pivotal in the discrepancy criteria used by many researchers to establish reading disability. (1987, pp.157–158)

Marie Clay's comment on the ability of Down's Syndrome children to acquire functional reading skills is reminiscent of my own clinical experiences in which I encountered many children with very low (measured)

intelligence who had acquired functional word and pseudoword decoding skills although their comprehension skills were typically marginal or worse (such children have often been called "hyperlexic"). The fact that these children had managed to master the alphabetic code while others I worked with had great difficulty doing so, despite having average or above average intelligence, was one of many reasons I came to have little faith in the use of intelligence tests to identify specific reading disability.

In the final section of Part I, Marie asserts that LD and low-achieving readers cannot be distinguished on the basis of neurological, etiological, psychological, or achievement criteria nor on the basis of their reading behaviors. I was especially impressed with the following comments:

Suppose I have three children learning to read. Child 1 makes typical, normal, usual, progress for age. Child 2 does not have a learning disorder, but because of sickness, poor school attendance, an emotional problem, non-active search to solve problems, or some such state or set of events, she or he falls further and further behind classmates, and all self-scheduled learning efforts proceed on false or unhelpful assumptions. Child 3 does have minimally organically-based brain dysfunctioning; and though teacher, programme, and child try to get progress in reading, every day's experiences add unhelpfully to the way knowledge is stored, or processes are carried out. Our best evidence shows that both Child 2 and 3 build a system of responses which does not work efficiently, and in particular does not continue to build itself with self-scheduled learning. In this model, the longer the child is in an inappropriate programme, the greater the number of times that child will have practiced inappropriate processing, day after day, year after year. Such children are building highly practiced inappropriate response systems. There is no way to press the exit button from the programme and clear the computer. Such children are *learning to be LD* (italics mine) with increasing severity as long as the inappropriate responding continues. The distinction between disability in the child and event-produced responses is irrelevant. Both have difficulties of a grand order, both have difficulties hidden from view and we do not have really effective instruction for either group. (1987, p. 160)

Of course, *learning to be learning disabled* is an elegant characterization of the type of negative learning that accrues in a child who has not acquired the foundational knowledge that would allow him or her to actively engage in the type of “self-scheduled” teaching and learning that enables that child to build upon the classroom instructional program and fill in the inevitable gaps inherent in all classroom programs, even those of very high quality. And, here Marie Clay makes an important point about the limitations of classroom instruction for the struggling reader:

Normal reading progress comes about because you have an active child, actively processing during opportunities provided by the programme with enough teacher attention to tip a slightly off-side response back into efficient functioning modes. Class instructional systems are not geared to responding to poor readers who have each in his or her idiosyncratic way wandered away from the programme’s sequence of process building. Educational systems with group instruction were not developed to deliver treatments that are (a) tightly controlled, (b) individually designed, and (c) contingently delivered. Yet the child cannot usually actively process his own way back to normal response patterns. He or she keeps trying, the teacher keeps trying, but the response system, like a tangled ball of twine, gets more knotted with more effort. (1987, p. 160)

The informed reader cannot help but recognize, in these assertions, the basic philosophy motivating the earliest iteration of Marie Clay’s Reading Recovery — specifically the assumed need to provide intense, highly individualized, and comprehensive instruction that is responsive to the struggling reader’s unique profile of strengths and weaknesses. Reading Recovery, as initially implemented, was and continues to be a one-to-one tutoring program on the grounds that this format facilitates the type of dynamic interaction between student and teacher that is necessary for detecting negative and ineffectual reading behaviors that impede progress and replacing them with positive and effectual reading behaviors that promote progress. This is not to suggest that all struggling readers will require one-to-one tutoring for effective remediation of their reading difficulties, because there is strong evidence that both small-group and classroom-level intervention can be effective remedial formats for many struggling readers, especially if intervention is imple-

mented early in the child’s educational history (e.g., see Scanlon et al., 2008, and Vellutino et al., 2008 for recent evidence). Neither is it to suggest that Reading Recovery’s influence is restricted to its one-to-one tutorial setting. Reading Recovery teachers have a long history of providing small-group intervention and classroom consultation services on behalf of struggling readers and reserve one-to-one tutoring for the most severely impaired readers (Emily Rodgers, personal communication 7/3/2010; see also Schwartz, Schmitt, & Lose, 2008). Nevertheless, because one of the major objectives of our first intervention study was to use remedial intervention as a “first-cut” diagnostic to help distinguish between experientially and constitutionally impaired readers, it seemed likely that children who might fit the latter description would likely be among the most severely impaired readers, and Marie Clay’s assertion that effective intervention will necessarily be “tightly controlled, individually designed, and contingently delivered” (1987, p. 160) convinced us that one-to-one tutoring of the type used in Reading Recovery would maximize our ability to distinguish between these two groups of impaired readers.

In Part II of Marie Clay’s article (*Learning Events as Causes*), she expands upon concerns she expressed in Part I about the differential diagnosis of the basic causes of reading difficulties, especially in older children. She makes special note of the circularity inherent in the common use of measures which, themselves, depend on some degree of proficiency with spoken or written language to help detect organically based reading disability:

There is no way the psychologist can take written language responses of the 10 year old to be indicative of the organic nature of the brain because that brain has been learning to make those responses in interaction with the teaching delivered for five years. Any unusual spelling patterns or confusions about letter names and sounds have long histories of muddled learning. Scores on tests which use language symbols, written or spoken, for stimulus or response, will be partially if not entirely event-produced. Alphabet sequences, letter recognition, spelling patterns, auditory memory for verbal material—many of the diagnostic tests used to test specific aspects of limited functioning are heavily influenced by instruction, and the child’s efforts to respond to instruction. This usually amounts to years of inappropriate learning. In this sense, the

children being tested have been *learning to be learning disabled* (italics mine) in terms of the diagnostic batteries. (1987, p. 163)

However, the most important message in this section of the article is that struggling readers *learn to be learning disabled because they are taught to be learning disabled* as a consequence of the fact that their instructional programs are not responsive to their individual needs:

There are quality differences among teachers, and programmes are fallible and fallibly delivered. Add to all that the ups and downs in a child's personal learning life there must always be hiccoughs in the learning sequences which children go through. Putting these things together it is reasonable to assume that high progress readers can overcome the conceptual, quality, and random aspects of programmes if they are actively working on building their own reading response systems independently, just as most children build their own oral language system. The low progress readers are not only victims of the conceptual bias of the programme, but also of the quality of delivery and of the events in their personal life histories. (Clay, 1987, pp.164–165)

The author amplifies this message in discussing two models of teaching. One model “seeks to work from a diagnosis of what is wrong—either where the faulty brain functioning occurs or what the child cannot do” (Clay, 1987, p. 165) and is based on two related assumptions: (a) “that there is more than one route to learning appropriate responding” (p. 165), and (b) “that some kind of responding is not essential for effective functioning” (p. 165). In contrast, the second model “assumes that effective reading, though complex, requires a standard set of patterned functioning, all components of which are necessary” and “works on the assumption that strengths must be enlisted to support whatever is difficult and whatever is difficult must be learned” (pp. 165–166); and further, that effective reading is “best understood by studying normal reading processing” (p. 166) and that “teaching ingenuity adapted to pupil peculiarities is called for” (p. 166), in order to “devise tasks which lead the particular individual with particular patterns of responding slowly and gradually from where he or she is towards the fully-operative model of normal reading behaviours which is the goal” (p. 166). The contrast in

the likely effects of these two models is summarized quite nicely in the following:

Teachers run the risk of *teaching the child to be learning disabled* (emphasis mine) when they design their lessons from their models of disability according to the first model. The child is likely to learn many items and responses relevant to reading but to be unable to orchestrate the act of reading.

Teaching interventions using the second model aim to develop response repertoires adequate for processing successive information in a text, and evaluating and monitoring in order to detect when messages are in error. (1987, p. 166)

I found myself in complete agreement with these contentions because I had long been frustrated by disturbing and misguided educational practices characterized by activities designed to improve functioning in basic processes such as visual and auditory perception, sensory integration, and sensory-motor integration, in the interest of increasing learning ability and ameliorating learning disability in reading and other areas. Such “process training” activities, as they were often called, were implemented as an integral part of the child's individualized educational plan (IEP), in the absence of any scientific evidence supporting their use (Kavale & Mattson, 1983; see also American Academy of Pediatrics et al., 2009; Vellutino et al., 2004 for more recent reviews). In fact, I wrote an article entitled “Has the Perceptual Deficit Hypothesis Led Us Astray?”—published a full decade before I received Marie Clay's article—expressing many of the same concerns she expressed in her article (Vellutino, Steger, Harding, Moyer, & Niles, 1977). I would like to say that such process training activities are no longer evident in our educational institutions, but, sadly, that is not the case and certainly was not the case when Marie Clay sent me her article. Nevertheless, the points she made in this section of the article were quite in keeping with our own view that, regardless of the nature and severity of the child's reading difficulties, the only viable way to remedy those difficulties is to provide that child with direct, individualized, comprehensive, and responsive instruction in reading, spelling, and writing and her *learning to be learning disabled* perspective reinforced our belief that we were on the right track.

In the last major section of Marie Clay's article (Part III: Some Implications for Research on Learning Dis-

ability), she makes a strong case for using a longitudinal approach to LD research in concert with comprehensive intervention, in the interest of distinguishing between experiential and constitutional causes of reading difficulties. She suggests that “the longitudinal approach has a contribution to make in this area, not because it will uncover variables that will predict reading difficulties, but because it offers the opportunity to describe the onset and progression of the condition, particularly the aspects that are learned in interaction with the task” (1987, p. 166–167). She also suggests that longitudinal study can help determine the relative importance of given components of literacy in individual children at different stages of their development in reading skills acquisition, for example fluency in word identification as compared with components of language comprehension (e.g., vocabulary and syntactic knowledge) as determinants of reading comprehension in more- versus less-advanced readers (examples are mine; see Vellutino, Tunmer, Jaccard, & Chen, 2007 for research evidence).

However, Clay points out that “the longitudinal approach will not of itself design or determine the parameters of the intervention delivery” (p. 167), which I took to mean that the nature and quality of the intervention will necessarily be determined by the knowledge, experience, and “ingenuity” of the interventionist. Accordingly, the article then focuses on Reading Recovery as a model of the type of intervention that holds promise for distinguishing between experiential and constitutional causes of reading difficulties. Thus, in commenting on Bryant and Bradley’s (1985) suggestion that training studies hold promise for detecting causal relationships, Clay asserted that

This may be possible in a simple task but simple causal hypotheses about the complex activity of reading or writing may not be revealed by such an approach. An early intervention programme may, however, assist the scientist studying LD to clean up his or her research samples. Reading Recovery has this potential. (1987, p. 169)

And, in a later section, she asserts

Reading Recovery is a programme which should clear out of the remedial education system all the children who do not learn to read for many event-produced reasons and all the children who have organically-based problems but who can be taught to achieve independent learning status in reading and writing despite this, leaving a small group of

children requiring specialist attention. I assume that the small group will still contain children whose problems are event-produced and children whose problems have organic origins. (1987, p. 169)

## Documentation for Marie Clay’s Suggested RTI Approach to Reading Disability Classification

It should be apparent to the reader that these latter comments embody what is currently described as the RTI approach to reading disability classification, and I believe that Marie Clay was the first reading researcher to suggest and use this approach, as I indicated earlier. More important is the model she provided for conducting reading disability research, which we took quite seriously insofar as it became the basic research design we used for our first intervention study (Vellutino et al., 1996) and for subsequent intervention studies (Scanlon et al., 2008; Vellutino et al., 2008). In fact, we gave her full credit for the ideas that convinced us to use this design (Vellutino et al., 1996). Because results from two of these studies were quite in keeping with Marie Clay’s *learning to be learning disabled* perspective while providing seminal documentation for the RTI approach to reading disability classification, I will briefly describe these studies and summarize their salient findings.

### *The Vellutino et al. 1996 Study*

In keeping with Marie Clay’s suggestions, our first intervention study used a longitudinal design that incorporated a major intervention component as a means of distinguishing between children whose reading difficulties were likely caused primarily by limitations in early literacy experiences and/or less than adequate literacy instruction, and those whose reading difficulties were caused primarily by basic deficits in cognitive abilities presumed to underlie reading ability. To accomplish this objective, we conducted a 5-year longitudinal study in which we periodically assessed struggling readers on measures of reading achievement and reading-related cognitive abilities from the beginning of kindergarten through the end of fourth grade, before and after they received remedial intervention. Of special interest was the question of whether the cognitive profiles of children who were found to be difficult to remediate would differ from those who were found to be readily remediated, relative to normally developing readers.

The majority of the struggling readers ( $n = 76$ ) were provided with daily one-to-one tutoring (30 minutes per session), using an approach to intervention, much like Reading Recovery, that we have come to call the Interactive Strategies Approach (IAS) (Scanlon, Anderson, & Sweeney, 2010; Vellutino & Scanlon, 2002). The rest ( $n = 42$ ) were provided with whatever remedial services were offered at their home schools and a random trials procedure was used to assign children to each of these conditions. Intervention was initiated in mid-first grade and was terminated at the end of first grade for children who were found to be readily remediated and in mid-second grade for children who were found to be more difficult to remediate. Thus, all of the more-difficult to remediate children were provided with two consecutive semesters of the intervention program. However, after one semester of

---

The data also suggested that RTI may be a more valid means of identifying specific reading disability than the more traditional IQ-achievement discrepancy, which is quite in keeping with Marie Clay's suggestion that the use of intelligence-based criteria for defining learning disabilities is not a defensible practice.

---

exposure to the program, all of the tutored children who received project-based intervention were rank ordered on the basis of measures of growth in reading during that semester. The reading growth measures on which the children were rank-ordered were regression slopes obtained by regressing scores on the Woodcock Reading Mastery Test-Revised (WRMT-R, Woodcock, 1987) Basic Skills Cluster on time in months from December of first grade to September of second grade. The children were thereafter separated into four (approximately equal groups designated as follows: Very Limited Growth (VLG), Limited Growth (LG), Good Growth (GG), and Very Good Growth (VGG). For purposes of comparison, two groups of normally developing readers were also identified in mid-first grade. One group consisted of children with average intelligence (AvIQNorm,  $n = 28$ ) and a second group consisted of children with above aver-

age intelligence (AbAvIQNorm,  $n = 37$ ). Reading growth in these children was also periodically assessed from the beginning of kindergarten through the end of fourth grade. In addition, all groups were compared on measures of intelligence and reading-related cognitive abilities in kindergarten, first, and third grade.

First note that the children in the tutored group generally performed better than the children in the school-based comparison group on reading outcome measures, suggesting that our intervention was reasonably effective. Our results also indicated that the four tutored groups did not differ on the measures of intelligence, nor did they differ from the average IQ normal readers on these measures. In contrast, the struggling readers who were found to be difficult to remediate differed significantly from the struggling readers who were found to be readily remediated on measures of language-based skills, especially phonological skills that are important for learning to read (e.g., knowledge of letter names and sounds, phoneme awareness, letter-sound decoding, verbal memory, name retrieval, etc.). In addition, the children who were readily remediated performed at levels comparable to those of the normal readers on these measures. At the same time, the normal reader groups did not differ on measures of word level skills. Thus, results from this study showed no strong or reliable relationship between measures of intelligence and measures of reading achievement, either among struggling readers who responded differentially to intensive intervention or among normally developing readers. The data also suggested that RTI may be a more valid means of identifying specific reading disability than the more traditional IQ-achievement discrepancy, which is quite in keeping with Marie Clay's suggestion that the use of intelligence-based criteria for defining learning disabilities is not a defensible practice (see also Vellutino, Scanlon, & Lyon, 2000, for a detailed commentary).

However, in order to provide more concrete evidence for the latter assertion, we performed several additional analyses that were not reported in the Vellutino et al. (1996) paper. In one such analysis we correlated measures of word identification, word attack, and comprehension skills with measures of verbal and nonverbal intelligence in both the tutored groups and the normal reader groups. Table 1 presents results from these analyses. The first finding of note is that there were very few statistically significant correlations between the various measures of intelligence and the various measures of reading achieve-



**Table 1. Correlation Coefficients Between IQ Scores on Reading Measures at Different Time Points for Normal and Tutored Readers in the Vellutino et al. (1996) Intervention Study**

IQ Score	Grade 1 Winter	Grade 1 Spring	Grade 2 Fall	Grade 2 Winter	Grade 2 Spring	Grade 3 Spring	Grade 4 Spring
<b>Word Identification and IQ</b>							
<b>Normal Readers<sup>a</sup></b>							
Verbal IQ							
Grade 1	.15	.13	.15	.14	.23	.18	.08
Grade 3	.25	.26	.35**	.37**	.42**	.32*	.32*
Performance IQ							
Grade 1	.03	.11	.07	.18	.11	.12	.21
Grade 3	.03	.10	.12	.20	.13	.18	.30*
<b>Tutored Readers<sup>b</sup></b>							
Verbal IQ							
Grade 1	.12	.12	.16	.05	.06	.06	.16
Grade 3	.24*	.23*	.29*	.16	.12	.17	.25*
Performance IQ							
Grade 1	.15	.23	.18	.00	.05	.05	.01
Grade 3	-.01	-.01	-.12	-.17	-.13	-.06	-.08
<b>Word Attack and IQ</b>							
<b>Normal Readers<sup>a</sup></b>							
Verbal IQ							
Grade 1	.16	.06	.06	.06	.06	.06	.04
Grade 3	.23	.16	.27	.31*	.3	.22	.25
Performance IQ							
Grade 1	-.00	-.00	.10	.04	.06	-.06	.15
Grade 3	.21	.03	.17	.19	.14	.15	.40**
<b>Tutored Readers<sup>b</sup></b>							
Verbal IQ							
Grade 1	.05	.08	-.02	-.06	-.02	-.08	-.01
Grade 3	.17	.22	.14	.09	.14	.06	.05
Performance IQ							
Grade 1	-.00	.19	.05	.09	-.09	-.18	-.05
Grade 3	.03	-.06	-.11	-.15	-.15	-.16	-.16
<b>Reading Comprehension and IQ</b>							
	<b>Grade 1</b>	<b>Grade 3</b>		<b>Grade 1</b>	<b>Grade 3</b>		
<b>Normal Readers<sup>a</sup></b>			<b>Tutored Readers<sup>b</sup></b>				
Verbal IQ			Verbal IQ				
Grade 1	.25*	.55**	Grade 1	.10	.11		
Grade 3	.30*	.56**	Grade 3	.19	.22		
Performance IQ			Performance IQ				
Grade 1	.15	.36**	Grade 1	.05	.09		
Grade 3	.19	.28*	Grade 3	-.09	.04		

Note: Raw scores for the Word Identification and Word Attack subtests are taken from the Woodcock Reading Mastery Test-Revised. The raw scores for the reading comprehension tests are taken from the Passage Comprehension Subtest of the Woodcock-Johnson Psychoeducational Battery. Because of larger sample sizes, correlations of comparable magnitudes may achieve statistical significance in the tutored readers but not in the normal readers.

<sup>a</sup>n range = 51 to 65; <sup>b</sup>n range = 69 to 76; \*p ≤ .05 (two-tailed tests); \*\*p ≤ .01 (two-tailed tests)

ment in either group. Moreover, those that were found to be statistically significant tended to be low and inconsistent in both groups. This was especially true in the case of the measures of word identification and word attack. This finding is important because it is the discrepancy between a measure of word level skills and a measure of intelligence that has been typically used to define specific reading disability.

The second finding of note is that correlations that were found to be statistically significant when word level skills were the dependent measures emerged almost exclusively in the case of those involving the test of verbal intelligence obtained in third grade. This finding is important because it suggests that observed correlations between measures of reading achievement and measures of intelligence could often be an artifact of experience in reading that affects the acquisition of knowledge and skills that might influence performance on a test of intelligence, for example language-based skills such as vocabulary knowledge as well as other skills that depend heavily on reading ability (Stanovich, 1986). It is also in keeping with Marie Clay's suggestion that diagnostic tests that are used to define specific reading disability are far too often confounded by reading difficulties that impair the acquisition of knowledge assessed by such tests.

Finally, the observed pattern of correlations suggests that the predictive validity of IQ scores is, at best, weak. More specifically, none of the correlations between measures of intelligence administered at the end of first grade and the various reading measures administered subsequently was statistically significant when the measures of word identification and word attack were the dependent measures. This pattern was observed in both the tutored children and the normal readers.

It should also be noted that none of the correlations that involved reading comprehension was statistically significant in the tutored group, whereas most of the correlations were statistically significant in the normal reader group when reading comprehension was the dependent measure. This disparity is undoubtedly due to the fact that many of the children in the tutored group had not yet acquired enough facility in word identification for comprehension processes to become fully operative in text processing. Also, it is not surprising to find positive and significant correlations between measures of intelligence and measures of reading achievement when reading achievement is defined as the ability to comprehend

the meanings of given texts in struggling readers who have acquired enough facility in word decoding and word identification for language comprehension processes and background knowledge to become fully operative. In fact, it is indisputable that knowledge and skills assessed to some degree by most intelligence tests are important for language and reading comprehension (e.g. vocabulary knowledge, use of inference, reasoning, etc.), and will undoubtedly contribute to individual differences on reading measures that entail such knowledge and skills, even in children who have adequate reading skills. This inference is consistent with the finding that observed correlations were larger and more consistent in the normal reader group when reading comprehension was the dependent measure. Yet, as I have just noted, specific reading disability has been conventionally defined as a significant deficiency in word identification that adversely affects reading comprehension, and has never been defined as a significant deficiency in reading comprehension in struggling readers who have acquired adequate facility in word identification.

It seems reasonable to conclude from these results that measures of intelligence do not predict either the short- or long-term effects of intervention in struggling readers, in terms of distinguishing between those who continue to be at risk and those who are no longer at risk following intervention. It also seems that a more direct and more effective approach would be to use measures of children's initial response to intervention along with measures of their ability to consolidate and maintain the gains they achieved as a result of the intervention as predictors of short- and long-term reading achievement. To assess this possibility, we conducted a series of hierarchical regression analyses using measures of RTI as predictors of subsequent reading achievement after controlling for IQ as well as for baseline differences and autocorrelation with the dependent measures. The RTI measures were: (a) simple gain scores on tests of basic word level skills that reflected growth in these skills from winter to spring of first grade (that is, after all children had received one school semester of remedial intervention); and (b) change scores on the same tests from spring of first grade to fall of second grade, which we used as operational measures of the child's ability to maintain gains facilitated by the intervention over the summer hiatus (henceforth called the "summer drop off").<sup>1</sup>

The dependent measures in the hierarchical regression analyses were Word Identification, Word Attack, and

Basic Skills Cluster (BSC) standard scores from the Woodcock Reading Mastery Test-Revised (Woodcock, 1987), and Woodcock-Johnson Passage Comprehension standard scores (Woodcock & Johnson, 1989). Each of these measures was obtained from all available tutored children in spring of second, third, and fourth grade. The primary independent measures used as predictors for each set of analyses were first-grade winter-to-spring gain scores and summer drop off change scores on given measures. To control for baseline differences as well as autocorrelation with the dependent measures, we entered the first-grade winter baseline scores for each predictor

---

It seems reasonable to conclude from these results that measures of intelligence do not predict either the short- or long-term effects of intervention in struggling readers, in terms of distinguishing between those who continue to be at risk and those who are no longer at risk following intervention. It also seems that a more direct and more effective approach would be to use measures of children's initial response to intervention along with measures of their ability to consolidate and maintain the gains they achieved as a result of the intervention as predictors of short- and long-term reading achievement.

---

into the predictive equations, prior to the gain scores and summer drop off scores, with the gain scores preceding the summer drop off scores. In the first set of analyses, the Word Identification standard scores were the dependent measures and the Word Attack standard scores were the predictors. In the second set of analyses, the Word Attack standard scores were the dependent measures and the Word Identification standard scores were the predictors. Each of these analyses was designed to assess the degree to which initial growth in one basic word level

skill would predict long-term performance in a related (basic) word level skill. In the third set of analyses, the BSC standard scores were used as both dependent measures and predictors, in order to assess the degree to which initial growth in basic word identification and word decoding skills would predict growth in the same skills. In the fourth and final analysis, the Passage Comprehension standard scores were the dependent measures and the BSC standard scores were the predictors. However, to assess the predictive validity of intelligence test scores in predicting reading achievement, the Weschler Intelligence Scale for Children-Revised (WISC-R, Wechsler, 1974) Full Scale IQs of the tutored children obtained in first grade were entered into each of the equations before any of the other predictors.

Results from each set of analyses are presented in Table 2. As can be seen, the first-grade Full Scale IQ scores contribute no significant variance in predicting performance on any of the dependent measures at the end of second, third, and fourth grade, as would be expected from the simple correlations reported in Table 1. In contrast, the first-grade gain scores and summer drop off scores contribute significant and unique variance in predicting performance on each of the dependent measures at all assessment periods, even after controlling for baseline differences on these predictors. The Word Attack gain scores and the summer drop off scores explained between 29% and 40% of unique variance on the Word Identification outcome measures at the end of second, third, and fourth grade; the Word Identification gain scores and summer drop off scores explained between 23% and 30% of unique variance on the Word Attack outcome measures at the end of these grades, and the BSC gain scores and summer drop off scores explained between 35% and 36% of unique variance on the BSC outcome measures over the same assessment periods. Finally, the first-grade BSC gain scores and summer drop off scores together explain 17% of unique variance on the Passage Comprehension outcome measure at the end of third grade. Note also that on all measures and at all assessment periods, the summer drop off scores accounted for significant unique variance, the amounts ranging from 3% to 12% (see O'Connor, Harty, & Fulmer, 2005 for similar results). These latter results suggest that a child's ability to consolidate and maintain intervention-based gains may be an important component of any RTI approach to reading disability classification. They also provide additional support for Marie Clay's contention that response to

**Table 2. R<sup>2</sup> and R<sup>2</sup> Change Estimates from Hierarchical Regression Analyses Documenting Increments of Variance in Predicting Performance of Tutored Children in the Vellutino et al. (1996) Study on the Word Identification, Word Attack, and Basic Skills Cluster at the End of Second, Third, and Fourth Grade, and Reading Comprehension at the End of Third Grade**

Measure	Second Grade		Third Grade		Fourth Grade	
	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change
<b>Word Identification</b>						
Full Scale IQ 1st Gr.	.002		.001		.014	
Word Attack 1st Gr. W	.064	.062*	.074	.073*	.059	.044
Word Attack 1st Gr. S-W	.350	.286**	.254	.180**	.295	.236**
Word Attack 2nd Gr. F-1st Gr. S	.469	.119**	.365	.111**	.380	.085**
<b>Word Attack</b>						
	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change
Full Scale IQ 1st Gr.	.009		.035		.003	
Word ID 1st Gr. W	.245	.236**	.208	.172**	.290	.287**
Word ID 1st Gr. S-W	.467	.222**	.397	.190**	.485	.195**
Word ID 2nd Gr. F-1st Gr. S	.499	.032**	.507	.110**	.523	.038*
<b>Basic Skills Cluster</b>						
	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change
Full Scale IQ 1st Gr.	.000		.001		.002	
Basic Skills 1st Gr. W	.379	.379**	.270	.269**	.290	.288**
Basic Skills 1st Gr. S-W	.685	.306**	.521	.251**	.581	.291**
Basic Skills 2nd Gr. F-1st Gr. S	.733	.048**	.628	.107**	.636	.056**
<b>Passage Comprehension</b>						
			R <sup>2</sup>	R <sup>2</sup> Change		
Full Scale IQ 1st Gr.			.003			
Basic Skills 1st Gr. W			.233	.230**		
Basic Skills 1st Gr. S-W			.365	.131**		
Basic Skills 2nd Gr. F-1st Gr. S			.402	.038**		

Note: All measures were standard scores. Gain scores were simple difference scores obtained by subtracting a child's winter baseline score on a given measure from that child's score on the same measure at the end of first grade. Summer drop off change scores were also simple difference scores obtained by subtracting a child's end-of-first grade score on a given measure from that child's score on the same measure at the beginning of second grade.

\* < .05; \*\* < .01

F = fall assessment; W = winter assessment; S = spring assessment; S-W = gain scores from winter to spring assessments; 2nd Gr. F -1st Gr. S = summer drop off change scores from end of first grade to beginning of second-grade assessments

intervention may ultimately prove to be a more valid criterion for diagnosing specific reading disability than the IQ-achievement discrepancy.

### *The Vellutino et al. 2008 Study*

This study was designed to develop a model for preventing long-term reading difficulties in kindergarten children at risk for early reading difficulties by providing them with supplementary intervention throughout the school year. A secondary objective was to further evaluate the efficacy of the RTI approach to diagnosing specific reading disability. To realize these objectives, children in schools that participated in the project were administered the WRMT-R Letter Identification subtest at the beginning of kindergarten, and all those scoring at or below the 30th percentile on this subtest were identified as being at risk for early reading difficulties. Half of the children in the at-risk group were randomly assigned to a project-based intervention group and the other half were assigned to a school-based comparison group. The children assigned to the intervention group received small-group supplemental instruction (3 children per group) for half-hour sessions twice weekly throughout kindergarten. Those assigned to the school-based comparison group received whatever remedial services were offered by their home schools. Literacy growth in all of the at-risk children was assessed throughout their kindergarten year. Literacy skills in all available at-risk children were again evaluated at the beginning of first grade and these children were dichotomized into two separate groups: one designated as Continued Risk (CR) and another designated as No-Longer-at Risk (NALR), based on a composite measure derived from scores on tests that were used to assess growth on the emergent literacy skills targeted in the kindergarten intervention program. For purposes of comparison, two groups of normally developing readers were also identified: one having average intelligence (AvIQNorm) and the other having above average intelligence (AbAvIQNorm). Half of the CR children received additional intervention from project staff throughout first grade. The intervention consisted of daily one-to-one tutoring (30-minute sessions) that was comprehensive, highly individualized, and responsive. The other half received school-based intervention (typically in small groups). Literacy growth among children in all of these groups was tracked through the end of third grade when the project ended. Measures of verbal and nonverbal intelligence from the Weschler Intelligence

Scale for Children-III (WISC-III, Wechsler, 1991) were administered in first and third grade, and measures of reading-related cognitive abilities were also administered in third grade.

Finally, in order to assess the predictive validity of RTI, measures of response to kindergarten intervention were used to classify children into the CR and NALR groups that were constituted at the beginning of first grade and measures of response to first-grade intervention, along with a measure of intelligence, were used to predict reading performance at the end of second and third grade. The intelligence measure used in the first-grade prediction analyses was an abbreviated version of the WISC-III (i.e., a composite created from the Vocabulary, Information, Block Design, and Picture Completion subtests).

Results from this study extend and complement results from the Vellutino et al. (1996) study. First, note that the kindergarten intervention group was found to perform better than the school-based comparison group on the kindergarten literacy outcome measures, which suggest that our kindergarten intervention was reasonably effective. In addition, approximately 84% (98/117) of the at-risk children became at least average level readers, after receiving either kindergarten intervention alone or both kindergarten and first-grade intervention combined. At the same time, 73% (72/98) of the children in this group received only kindergarten intervention, and all of these children continued to perform at least within the average range for the remainder of the project. Moreover, at-risk children who received project-based supplementary intervention in kindergarten performed significantly better than at-risk children who received no supplementary intervention in kindergarten on the measures of emergent literacy skills administered at the end of kindergarten and the beginning of first grade. In addition, the NALR children performed significantly better than the CR children (and close to the levels of the normal readers) on measures of reading-related language abilities included in the cognitive battery administered in third grade. However, these two groups were not significantly different on measures of nonverbal intelligence, in accord with results obtained in the Vellutino et al. (1996) intervention study.

These findings suggest that early and long-term reading difficulties can be prevented in the large majority of children who may be at risk for such difficulties if these children are identified at the beginning of kindergarten (if not earlier) and provided with the supplementary

intervention needed to correct deficiencies in foundational literacy skills. They also provide additional support for our contention that early reading difficulties in most children are caused by experiential and instructional deficits rather than biologically based cognitive deficits, which, of course accords quite well with Marie Clay's *learning to be learning disabled* perspective.

Another interesting finding in this study is that response to kindergarten intervention, defined in terms of gain scores on measures of emergent literacy skills, predicted membership in the CR group at the beginning of first grade with 90% accuracy and membership in the NALR group with 87% accuracy. Moreover, when we added a measure of change over the summer to the prediction equation (the summer drop off), we increased predictive accuracy to 95% for CR status and 96% for NALR status. These results essentially replicate results obtained in the Vellutino et al. (1996) study (as reported above). However, of special interest are results obtained in the CR children ( $n = 45$ ) who required and received additional and more intensive intervention (one-to-one tutoring) throughout first grade. Fifty eight percent (26/45) of these children performed at least in the average range on all reading outcome measures at the end of first, second, and third grade and at levels that approximated those of children in the NALR and normal reader groups at all these time periods. In contrast, 42% (19/45) of the children in this group performed below average on the reading outcome measures at the end of second and third grade, despite the fact that they performed in the low average to average ranges on these measures at the end of first grade. Since the children in both of these groups received project-based intervention in both kindergarten and first grade, those in the latter group were labeled Difficult-to-Remediate (DR) while those in the former group were labeled Less-Difficult-to-Remediate (LDR).

In order to compare the predictive validity of measures of intelligence with measures of RTI, in forecasting long-term reading achievement in the DR and LDR groups, hierarchical regression analyses were used to predict reading performance in these children at the end of second and third grade using both first grade IQ scores and measures of response to first-grade intervention as predictors. More specifically, the predictors used in these analyses were: (a) the Full Scale IQ from the abbreviated version of the WISC-III administered in first grade; (b)

beginning to end of first-grade gain scores on the Word Identification, Word Attack, and BSC of the WRMT-R, representing response to intervention provided throughout first grade; and (c) baseline scores on these measures to control for variability at the beginning of first grade and autocorrelation with the reading outcome measures. All three of the WRMT-R measures were also used as dependent measures in these analyses, along with the Passage Comprehension subtest of the WRMT-R. In each regression equation, the Full Scale IQ was always entered first; the baseline scores were entered next, followed by the RTI gain scores.

Results from these analyses are presented in Table 3. The first finding of note is that the first-grade Full Scale IQ accounted for no significant variance on any of the reading outcome measures administered at the end of second and third grade<sup>2</sup>, thus replicating results reported for similar analyses using data from the Vellutino et al. (1996) intervention study (as reported above). The second is that none of the first-grade baseline scores accounted for significant variance on the reading measures, which, of course, indicates that the children in the DR and LDR groups had comparable levels of reading skill at the beginning of first grade. But, of special significance is the finding that the first-grade RTI gain scores accounted for significant variance on all of the reading outcome measures administered at the end of second and third grade. The Word Attack gain scores accounted for 22% and 14% of the variance on the Word Identification subtest administered at the end of second and third grade (respectively); the Word Identification gain scores accounted for 44% and 54% of the variance on the Word Attack subtest at the same assessment periods; the BSC gain scores accounted for 57% and 54% of the variance on the same measure administered at the end of second and third grade; and the BSC gain scores accounted for 28% of the variance on the Passage Comprehension subtest administered at the end of third grade.

These results are consistent with results obtained in the Vellutino et al. (1996) study (as reported above) and further undermine the use of intelligence tests to diagnose specific reading disability. The combined findings provide additional evidence that IQ scores do not predict responsiveness to remedial intervention, whereas more-direct measures of initial response to intervention strongly predict such effects.

**Response to Intervention: Current Trends and Cautions**

Results from the two intervention studies summarized above, together, provide strong support for the major assertions constituting Marie Clay's *learning to be learning disabled* perspective. More specifically, our findings are consistent with her belief that reading difficulties in most struggling readers are caused by "events" in the child's

personal and educational histories rather than organically based cognitive deficits. The data are also consistent with her belief that response to intervention would ultimately prove to be a more effective and more valid means of distinguishing between "event-produced" and organic causes of reading difficulties than are psychometric/exclusionary approaches having the IQ-achievement discrepancy as the

central defining criterion. In both studies, the large majority of at-risk or struggling readers became at least average-level readers following intervention, thus obviating LD classification in these children, many of whom would have undoubtedly satisfied criteria for such classification. And, in both studies, IQ scores did not distinguish between struggling readers whose reading problems were difficult to remediate and those whose reading problems were less difficult to remediate, nor did they distinguish between struggling readers and normally developing readers of average intelligence. Moreover, they did not predict either the short- or long-term effects of intervention, in terms of levels of performance on reading outcome measures administered shortly after or long after intervention was discontinued. In contrast, measures assessing a child's response to intervention (i.e. gain scores and summer drop off scores) were shown to be significant predictors of both the short- and long-term effects of intervention.

Finally, our findings are quite in keeping with Marie Clay's assertion that reading instruction needs to be tightly controlled, individually designed, and contingently delivered in order for it to be responsive to a struggling reader's individual needs and capitalize on both his/her strengths and weakness. More specifically, the first-grade struggling readers in our first intervention study (Vellutino et al., 1996) received one-

**Table 3. R<sup>2</sup> and R<sup>2</sup> Change Estimates from Hierarchical Regression Analyses Documenting Increments of Variance in Predicting Performance on the Word Identification, Word Attack, and Basic Skills Cluster at the End of Second and Third Grade at the End of Third Grade for DR and LDR Readers Identified in the Kindergarten and First-Grade Intervention Study**

Measure	Second Grade		Third Grade	
	R <sup>2</sup>	R <sup>2</sup> Change	R <sup>2</sup>	R <sup>2</sup> Change
<b>Word Identification</b>				
Full Scale IQ 1st Gr.	.015		.063	
Word Attack 1st Gr. S	.054	.039	.127	.064
Word Attack 1st Gr. J-S	.278	.224**	.265	.138**
<b>Word Attack</b>				
Full Scale IQ 1st Gr.	.042		.023	
Word ID 1st Gr. S	.042	.000	.062	.000
Word ID 1st Gr. J-S	.481	.439**	.597	.535**
<b>Basic Skills Cluster</b>				
Full Scale IQ 1st Gr.	.028		.064	
Basic Skills 1st Gr. S	.037	.009	.075	.011
Basic Skills 1st Gr. J-S	.607	.571**	.610	.535**
<b>WIAT Reading Comprehension</b>				
Full Scale IQ 1st Gr.			.010	
Basic Skills 1st Gr. S			.028	.018
Basic Skills 1st Gr. J-S			.304	.276**

Note: All measures were standard scores. Gain scores were simple difference scores obtained by subtracting a child's September base-line score on a given measure from that child's score on the end-of-first-grade score on the same measure.

\*\* < .01

S = September; J-S = gain scores from September to June; WIAT = Wechsler Individual Achievement Test (Wechsler, 1992)

to-one remedial instruction that was much like Reading Recovery (Vellutino & Scanlon, 2002) and they, on average, performed better on reading outcome measures than did children in comparison groups who received school-based remediation, which was typically small-group instruction (see Schwartz et al., 2008 for similar results). Similarly, in our second intervention study (Vellutino et al., 2008), kindergarten at-risk children who received small-group remediation throughout the year performed better on measures of emergent literacy than did children who received “business-as-usual” classroom instruction. Thus, our findings accord quite well with Marie Clay’s predictions and with her perspective in general.

It should be clear that Marie Clay’s intervention-based approach to understanding the causes and correlates of early reading difficulties, as implemented in her Reading Recovery program, was indeed a seminal and important contribution to literacy theory and practice and stimulated a plethora of intervention studies, beginning with our own, that provided significant documentation for RTI approaches to reading disability classification in lieu of traditional IQ-based psychometric approaches (Bradley, Danielson, & Hallahan, 2002). That RTI approaches have begun to receive widespread acceptance among policymakers and educators here in the United States is evidenced in the most recent revision of the Individuals with Disabilities Education Act (2004), which allows state and local education agencies to use response to “evidence-based” intervention to identify children with learning disabilities, though it continues to allow conventional IQ-based approaches to identification as well. Yet, despite the promise of the RTI concept for changing the landscape in educational practice on behalf of struggling learners, I am inclined to believe that Marie Clay would be disappointed in current trends in the implementation of RTI approaches to servicing children at risk for long-term learning difficulties. I say this because the RTI approaches currently advocated by researchers or implemented by practitioners tend to emphasize the utility and effectiveness of such approaches in identifying those who may be afflicted with organically based learning disabilities rather than their utility and effectiveness in preventing learning difficulties in those at risk for such difficulties and in reducing the incidence of learning disabilities by diversifying, individualizing, and improving the instruction we provide these children. Space does not permit a full-fledged discussion of this issue, but my colleague Peter Johnston has written an excellent and

comprehensive paper discussing the issue and I will limit my discussion to a few of the major concerns he expresses in his paper.

Johnston’s central concern is that RTI, as it is currently being conceptualized, far too often “emphasizes the need to identify individuals with disabilities rather than emphasizing reducing the need for such identification” (Johnston, in press, p. 3). In expanding on this theme, he makes essentially the same distinction between alternative RTI objectives that I have made above and asserts that current implementations of RTI tend to conceptualize it as a “measurement problem,” the solution of which inheres in developing “standard protocols” for both assessment and intervention that would be used

---

Despite the promise of the RTI concept for changing the landscape in educational practice on behalf of struggling learners, I am inclined to believe that Marie Clay would be disappointed in current trends in the implementation of RTI approaches to servicing children at risk for long-term learning difficulties.

---

to maximize the probability of identifying learning disabilities of constitutional origin rather than conceptualize it, more productively, as an “instructional problem” that focuses on the need to develop the educational expertise, a flexible set of intervention models, and, ultimately, a qualified cadre of educational personnel of the type required to service the individual needs of the heterogeneous population of struggling learners teachers are asked to educate. Current RTI-based approaches to LD classification conceptualize the enterprise as a three-stage process, commonly called the “three-tier model,” whereby intervention is initially implemented in the child’s classroom through (a) adjustments in the instructional program and progress monitoring for all children, followed by (b) small-group intervention and progress monitoring for those who continue having difficulties, and (c) more-intensive and more-individualized intervention (i.e., smaller groups or one-to-one tutoring), for those who



still have significant learning difficulties, despite having been exposed to two previous waves of intervention. Such a three-stage process is a reasonable approach to identifying children in need of special services and developing instructional formats and procedures that are responsive to their individual needs, if these are the inherent goals in implementing the process, and Johnston has no quarrel with this approach as a structural model. What does concern him is that most three-tier models being advocated by researchers or implemented in practice incorporate standard protocol formats for both instruction and progress monitoring and such formats, he points out, rest on a faulty set of assumptions.

For example, it is commonly assumed that standard protocols have wide applicability regardless of a child's age, unique abilities, individual learning profile, or instructional history and regardless of the teacher's experience and level of expertise. A related assumption attributes successful remediation to the materials and procedures specified in standardized assessment and treatment protocols and/or increased "dosage" of scripted treatments rather than the expertise and good efforts of a responsive teacher. Conversely, unsuccessful remediation is often attributed to a "fixed trait" in the child's make-up rendering him/her what has been variously called "a treatment resister" (Torgesen, 2000) or a "chronic non-responder" (Fuchs, Stecker, & Fuchs, 2008). Finally, Johnston eschews common applications of RTI that "assume a narrow, behavioristic view of literacy, focusing centrally on accurate and rapid identification of words, rather than on literacy more broadly defined" (in press, p. 12).

Johnston then goes on to present empirical and logical justification for these concerns, after which he focuses on the merits of conceptualizing RTI as a preventative and instructional enterprise rather than an identification and measurement enterprise. His central argument throughout this section is that the success of the RTI paradigm shift hinges greatly on embracing the former rather than the latter option, and ultimately, on developing the expertise of educational personnel responsible for implementing RTI approaches:

If the emphasis is put on instruction, then evidence that the child is not learning adequately is primarily evidence that instruction is not yet appropriate and needs to be further optimized. It is evidence of a need for greater instructional expertise and perhaps a reduced teacher-student ratio. (in press, p. 15–16)

He adds: "Teacher expertise is the most important factor in improving children's learning, and children experiencing the most difficulty should have the most expert teachers" (p. 16), which is, far too often, simply not the case. Johnston provides empirical and logical justification for these assertions and cites Reading Recovery and our Interactive Strategies Approach (ISA) (Scanlon et al., 2010; Vellutino & Scanlon, 2002) as cardinal examples of the types of responsive instruction necessary to make RTI a viable enterprise. Reading Recovery and the ISA have both been shown to be effective with K–3 children using both one-to-one and/or small-group formats (Schwartz et al., 2008; Vellutino et al., 1996, 2008) and the ISA has recently been demonstrated to be an effective intervention in kindergarten classrooms (Scanlon, et al., 2008; see also Gelzheiser, Scanlon, Vellutino, Hallgren-Flynn, & Schatschneider (in press) for documentation for an intermediate grade version of ISA called ISA-Extended).

Johnston goes on to contrast the identification-measurement and the prevention-instructional approaches to RTI in terms of both assessment and instructional differences between the two in a way that makes the pitfalls of the former approach rather obvious. He concludes his paper with the following cogent assertions among others:

The current identification-measurement approach to RTI may lead to confident classification of children with SLD in literacy; however, focusing on classification will not necessarily lead to a real reduction in the number of children with reading difficulties. To persist with this path requires a belief not only that a substantial group of children have permanent intrinsic traits (SLDs) that need identification, but that, having identified those traits, we have instruction that specifically addresses the handicap. There is no evidence of this. (in press, p. 31)

Johnston thereafter concludes:

If we want to capitalize on the promise of RTI, we will focus on prevention-instruction models, recognizing the complexity of literacy, its teaching and its learning, and centralizing the ongoing development of teacher expertise. None of this can be purchased in canned packages. (in press, p. 33)

I think that Marie Clay would strongly agree with these assertions, and so do I.

## Summary and Conclusions

To conclude, I simply wish to say that Marie Clay was right and her learning to be learning disabled perspective, as articulated in her excellent paper, was prophetic. Research in the study of learning disabilities and special educational practice are both in the throes of a paradigm shift and Marie Clay's work was at the root of this shift. Her Reading Recovery program was clearly the prototype for RTI approaches to identifying children at risk for long-term reading difficulties, some of whom might well be impaired by organically based cognitive deficits as she herself pointed out. Ironically, she has not been credited with her enormous and seminal contribution to this paradigm shift, as I indicated earlier. I have seen no references to her work in the RTI literature I have read, and as pointed out by Peter Johnston (Johnston, in press), Reading Recovery is not mentioned as a potentially effective RTI intervention on the federal government's What Works Clearinghouse website (<http://ies.ed.gov/ncee/wwc/>), despite the fact that it is the only intervention listed on the website as an effective intervention for improving reading ability in struggling readers. However, I have, in previous writings, given her the credit she deserves for her contribution to RTI research and practice, and the personal views I have expressed in this paper along with the data I have presented constitute a reaffirmation of this sentiment and empirical support for her perspective.

### Endnotes

<sup>1</sup> The end-of-first-grade gain scores were simple difference scores obtained by subtracting a child's winter base-line score on a given measure from the score obtained by that child on the same measure at the end of first grade. The summer drop off change scores were also simple difference scores obtained by subtracting a child's end-of-first grade score on a given measure from that child's score on the same measure at the beginning of second grade.

<sup>2</sup> Note that the WISC III subtest composite administered in first grade did not reliably differentiate the children in the DR and LDR groups, nor did it differentiate these groups from the NLAR and AvIQNorm groups (data not shown). In addition, this measure was not significantly correlated with any of the reading measures administered to the children in the DR and LDR groups (data not shown). The Verbal IQ administered in third grade did differentiate the DR group from the LDR, NLAR and AvIQNorm groups (data not shown), but this was likely due to the negative effects of prolonged reading difficulties (see Vellutino et al., 2003 for similar results). The Performance IQ (a measure of nonverbal intelligence) did not, however, differentiate these groups.

### References

- American Academy of Pediatrics, Section on Ophthalmology, Council on Children with Disabilities, American Academy of Ophthalmology, American Association for Pediatric Ophthalmology and Strabismus, and American Association for Certified Orthoptists (2009). Learning disabilities, dyslexia, and vision. *Pediatrics*, *124*, 837–844.
- Bradley, R., Danielson, L., & Hallahan, R. (Eds.). (2002). *Identifying learning disabilities: Research to practice*. Mahwah, NJ: Erlbaum.
- Bryant, P., & Bradley, L. (1985). *Children's reading problems*. Oxford, United Kingdom: Blackwell.
- Clay, M. M. (1987). Learning to be learning disabled. *New Zealand Journal of Educational Studies*, *22*(2), 155–173.
- Fuchs, D., Stecker, P. M., & Fuchs, L. S. (2008). Tier 3: Why special education must be the most intensive tier in a standards-driven, No Child Left Behind world. In D. Fuchs, L. S. Fuchs, & S. Vaughn (Eds.), *Response to Intervention: A framework for reading educators* (pp. 71–104). Newark, DE: International Reading Association.
- Gelzheiser, L. M., Scanlon, D. M., Vellutino, F. R., Hallgren-Flynn, L., & Schatschneider, C. (in press). Effects of the Interactive Strategies Approach-Extended: A responsive and comprehensive intervention for intermediate grade struggling readers. *Elementary School Journal*.
- Henderson, E. H. (1990). *Teaching spelling*. Boston, MA: Houghton-Mifflin.
- Individuals with Disabilities Education Improvement Act (2004). Public Law 108-446 (20 U.S.C. 1400 *et seq.*).
- Johnston, P. H. (in press). Response to intervention in literacy. *Elementary School Journal*.
- Kavale, K., & Mattson, P. D. (1983). One jumped off the balance beam: Meta-analysis of perceptual-motor training. *Journal of Learning Disabilities*, *16*(3), 163–173.
- O'Connor, R. E., Harty, K. R., & Fulmer, D. (2005). Tiers of intervention in kindergarten through third grade. *Journal of Learning Disabilities*, *38*(6), 532–538.
- Scanlon, D. M., Anderson, K. L., & Sweeney, J. M. (2010). *Early intervention for reading difficulties: The interactive strategies approach*. New York: Guilford Press.
- Scanlon, D. M., Gelzheiser, L. M., Vellutino, F. R., Schatschneider, C., & Sweeney, J. M. (2008). Reducing the incidence of early reading difficulties: Professional development for classroom teachers vs. direct interventions for children. *Learning and Individual Differences*, *18*, 346–359.
- Schwartz, R. M., Schmitt, M. C., & Lose, M. K. (2008, December). *The effect of teacher-student ratio on early intervention outcomes*. Paper presented at the National Reading Conference, Orlando, FL.

- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly, 21*, 360–407.
- Torgesen, J. K. (2000). Individual differences in response to early interventions in reading: The lingering problem of treatment resisters. *Learning Disabilities Research and Practice, 15*(1), 55–64.
- Torgesen, J. K., Rose, E., Lindamood, P., Conway, T., & Garvan, C. (1999). Preventing reading failure in young children with phonological processing disabilities: Group and individual responses to instruction. *Journal of Educational Psychology, 91*, 579–594.
- United States Office of Education (1977). Assistance for states for education of handicapped children: Procedures for evaluating specific learning disabilities. *Federal Register, 42*(50), 65082–65085.
- Vellutino, F. R., Fletcher, J. M., Snowling, M. J., & Scanlon, D. M. (2004). Specific reading disability (dyslexia): What have we learned in the past four decades? *Journal of Child Psychology and Psychiatry, 45*(1), 2–40.
- Vellutino, F. R., & Scanlon, D. M. (2002). The interactive strategies approach to reading intervention. *Contemporary Educational Psychology, 27*(4), 573–635.
- Vellutino, F. R., Scanlon, D. M., & Lyon, G. R. (2000). Differentiating between difficult-to-remediate and readily remediated poor readers: More evidence against the IQ-Achievement discrepancy definition of reading disability. *Journal of Learning Disabilities, 33*(3), 223–238.
- Vellutino, F. R., Scanlon, D. M., Sipay, E. R., Small, S. G., Pratt, A., Chen, R., & Denckla, M. B. (1996). Cognitive profiles of difficult-to-remediate and readily remediated poor readers: Early intervention as a vehicle for distinguishing between cognitive and experiential deficits as basic causes of specific reading disability. *Journal of Educational Psychology, 88*(4), 601–638.
- Vellutino, F. R., Scanlon, D. M., Zhang, H., & Schatschneider, C. (2008). Using response to kindergarten and first grade intervention to identify child at risk for long-term reading difficulties. *Reading and Writing: An Interdisciplinary Journal, 21*, 437–480.
- Vellutino, F. R., Steger, B. M., Moyer, S. C., Harding, C. J., & Niles, J. A. (1977). Has the perceptual deficit hypothesis led us astray?: An examination of current conceptualizations in the assessment and treatment of exceptional children. *Journal of Learning Disabilities, 10*, 375–385.
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a component skills model of reading development. *Scientific Studies of Reading, 11*(1), 3–32.
- Wechsler, D. (1974). *Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.



## About the Author

Dr. Frank Vellutino is professor of psychology at The University at Albany, the State University of New York. He currently holds joint faculty appointments in the Department of Psychology (Cognitive Psychology Program) and the Department of Educational and Counseling Psychology. He is also director of the University's Child Research and Study Center, a research and student training center. Dr. Vellutino's research has been concerned with reading development, the cognitive underpinnings of reading, and the relationship between reading difficulties and various aspects of language and other cognitive functions. His most recent studies have been concerned with early identification of children at risk for reading difficulties and early intervention on behalf of such children, in the interest of developing assessment and remedial procedures for preventing long-term reading difficulties.

- Wechsler, D. (1991). *Wechsler Intelligence Scale for Children-III*. New York: Psychological Corporation.
- Wechsler, D. (1992). *Wechsler Individual Achievement Test*. New York: Psychological Corporation.
- Woodcock, R. W. (1987). *Woodcock Reading Mastery Tests-Revised*. Circle Pines, MN: American Guidance Services.
- Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psychoeducational Battery-Revised*. Allen, TX: DLM Teaching Resources.

## Acknowledgements

The research discussed in this paper was supported by grants from the National Institute of Child Health and Human Development (NICHD) to Vellutino and his colleagues. The data for the intervention study reported in Vellutino et al. (1996) were collected as part of a project conducted under the auspices of a special Center grant (P50HD25806) awarded to the Kennedy Krieger Institute by NICHD. The data for the Vellutino et al (2008) intervention study were collected under the auspices of NICHD grant R01HD34598. The author wishes to express his sincere gratitude to the teachers, students, and secretarial and administrative staff in participating schools. He is also grateful to the intervention teachers and data collection personnel who participated in these projects. Finally, special thanks and gratitude are extended to my colleagues Donna Scanlon who was my co-investigator in both of the studies reported in this paper and to Sheila Small, Diane Fanuele, and Joan Sweeney who assisted immensely in the implementation of each study.